Ocean Engineering 48 (2012) 59–68

Contents lists available at SciVerse ScienceDirect

Ocean Engineering

journal homepage: www.elsevier.com/locate/oceaneng

Vision-based object detection and tracking for autonomous navigation of underwater robots

Donghwa Lee [a], Gonyop Kim [b], Donghoon Kim [b], Hyun Myung [a,b,*], Hyun-Taek Choi [c]

[a] Department of Civil and Environmental Engineering, KAIST 291, Daehak-ro, Yuseong-gu, Daejeon 305-701, Republic of Korea
[b] Robotics Program, KAIST 291, Daehak-ro, Yuseong-gu, Daejeon 305-701, Republic of Korea
[c] KORDI 171, Jang-dong, Yuseong-gu, Daejeon 305-343, Republic of Korea

ARTICLE INFO

Article history:
Received 21 April 2011
Accepted 7 April 2012
Editor-in-Chief: A.I. Incecik
Available online 1 May 2012

Keywords:
Underwater robot
Underwater vision
Underwater image restoration
Object detection
Object tracking

ABSTRACT

Underwater robots have been an emerging research area being at the intersection of the field of robotics and oceanic engineering. Their applications include environmental monitoring, oceanographic mapping, and infrastructure inspections in deep sea. In performing these tasks, the ability of autonomous navigation is the key to a success, especially with the limited communications in underwater environments. Considering the highly dynamic and three-dimensional environments, the autonomous navigation technologies including path planning and tracking have been one of the interesting but challenging tasks in the field of study. Cameras have not been at the center of attention as an underwater sensor due to the limited detection ranges and the poor visibility. Use of visual data from cameras, however, is still an attractive method for underwater sensing and it is especially effective in the close range detections. In this paper, the vision-based object detection and tracking techniques for underwater robots have been studied in depth. In order to overcome the limitations of cameras and to make use of the full advantages of image data, a number of approaches have been tested. The topics include color restoration algorithm for the degraded underwater images, detection and tracking methods for underwater target objects. The feasibilities of the proposed algorithms have been demonstrated in the experiments with an underwater robot platform and the results have been analyzed both qualitatively and quantitatively.

1. Introduction

Underwater robots, also known as autonomous underwater vehicles (AUVs), have been one of the main research areas in oceanic engineering for the past few decades. Their emerging applications include infrastructure inspections in deep sea, environmental monitoring, and oceanographic mapping. AUVs are considered as a cost-effective alternative to remotely operated vehicles (ROVs) where the untetheredness of AUVs gives higher maneuverability compared to the ROVs.

Considering the highly dynamic and three-dimensional underwater environments, autonomous navigation of underwater robots has been one of the interesting but challenging tasks in the field of study. For the majority of the applications, path planning and tracking are essential for the successful performance of an underwater robot. Path following to a docking station and area coverage for monitoring purposes are a few examples. Considerable amount of research has been conducted for the navigation of underwater robots and a number of techniques have been explored (Kim and Eustice, 2009a,b; Caffaz et al., 2010; Psarros et al., 2010; Marani and Choi, 2010; Balasuriya et al., 1997; Feezor et al., 2001; Corke et al., 2007).

The use of vision data from cameras is one of the useful methods for autonomous underwater navigations. Humans and most of the autonomous land and space systems use visions for navigation as it carries much information on the environment (Papanikolopoulos and Khosla, 1993). Cameras are considered as a cheaper solution with richer information compared to other underwater sensors such as acoustic sensors. Although the limited range of vision sensors and the large portion of featureless regions in underwater environments have been the downsides, the visual data can still play an important role in the underwater navigation, especially for close object and range detections.

Underwater images are known to be degraded due to a number of factors such as turbidity, floating particles, and light attenuation in the medium. Researchers have been trying to restore or enhance the degraded image qualities using various methods. Schettini and Corchs (2010) reviewed some of the most recent methods of underwater image processing. They nicely defined the two different points of view in underwater image processing as the image

* Corresponding author at: Department of Civil and Environmental Engineering, KAIST 291, Daehak-ro, Yuseong-gu, Daejeon 305-701, Republic of Korea.
Tel.: +82 42 350 3630; fax: +82 42 350 4540.
E-mail address: hmyung@kaist.ac.kr (H. Myung).

0029-8018/$ - see front matter © 2012 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.oceaneng.2012.04.006

restoration techniques and image enhancement methods. Trucco and Olmos-Antillon (2006) proposed a self-tuning restoration algorithm based on the simplified version of the Jaffe–McGlamery model. Although they showed the performance improvements in the image classification, the processing for a single 320×240 image took up to 3.8 s. Yamashita et al. (2007) introduced a color registration method with the consideration of light attenuation in water column. They visually recovered the degraded color information, however, they showed neither quantitative nor qualitative results.

Due to the poor visibility and the undesired effects such as aquatic snow, the object detection using cameras in underwater environments is considered as a challenging task. Yu et al. (2001) successfully demonstrated the detection of artificial underwater landmarks (AULs) using color extraction method. They also investigated a number of colors that are highly observable in underwater environments. They built the artificial landmarks using these colors and used the landmarks for the localization of their AUV.

More effort has been made for the better detection of the targets by adopting computer vision techniques. Dudek et al. (2005) designed color correction filter for reef detection using their amphibious legged robot, AQUA. They also designed visual servo control system that detects human guidance using vision sensors. As an extension of this work, Sattar and Dudek (2009) recently showed their work in underwater visual servo control. They demonstrated the performance of the underwater object detection using the ensemble tracker. The ensemble tracker is a combination of the weighted Haar-like detectors and it showed better performance compared to the color blob tracking method.

A number of researchers have proved the effectiveness of the visual guidance for the docking and homing of AUVs. Lee et al. (2003) designed the target object which guided the AUV to the docking position. The target is an LED ring with five large lights and it is detected by a single camera.

Some researchers have emphasized the advantages of the passive visual targets compared to the active targets. Negre et al. (2008) proposed a vision-based target detection method using passive targets. They showed their unique design of the underwater target called Self-Similar Landmark (SSL) which enables range estimations. Maire et al. (2009) also proposed a docking system with passive targets. They used Haar rectangular features to detect the encoded poles in the docking station.

Balasuriya et al. (1997) have developed an autonomous navigation system for underwater cable tracking using a single camera and an acoustic sensor. They used the two-dimensional data from the camera and the other dimension from the acoustic sensor. They used Laplacian of Gaussian (LoG) operator in image processing to minimize the effect of undesirable features. The three hierarchical controllers have been designed and the sequential control of surge, sway, and yaw of an AUV has been achieved while aligning to the cable. Balasuriya and Ura (2002) studied further in this area and proposed their algorithm of cable tracking using a priori map of the cable. They rather focused on the navigation of the robot using the given map when the cable features are invisible.

In this paper, autonomous navigation techniques for underwater robots using visually detected targets have been studied. The topics include real-time underwater image restoration methods, detection, and tracking of artificial landmarks. The main focus is to evaluate the effectiveness of the vision-based navigation methods in the underwater environments and to improve the performances using various techniques.

The contributions of this work include (i) testing feasibilities of various vision processing algorithms with underwater images and (ii) implementing and evaluating the proposed underwater color restoration method in real time.

The remainder of this paper is organized as follows. Section 2 introduces the underwater image restoration method that recovers the camera distortions and the color degradations. Vision-based object detection and tracking are explained in detail in Section 3. In Section 4, details of the experiments for the verification of the proposed algorithms as well as their results are described. And the conclusions are addressed in Section 5.

2. Underwater image color restoration

This section deals with the color restoration algorithm that recovers the degraded quality of the images acquired in underwater environments. Underwater images are essentially characterized by their poor visibility because the light is exponentially attenuated as it travels in the water and the resulting scenes are poorly contrasted and hazy (Schettini and Corchs, 2010). Recovery of the degraded color information should improve the results of the vision processing algorithms. For example, restoration of the color information will improve the object detection results when the objects are known or the pictures of them are given in the air. If the restored images express the interest objects in more realistic, or more like the ones in the air, many vision-based algorithms should take advantages of them, especially when color information is critical.

The underwater image recovery methods can be classified into two types according to two different points of view: image restoration and image enhancement. The image restoration aims to recover a degraded image quantitatively using a degradation model whereas the image enhancement uses qualitative subjective criteria to produce a more visually pleasing image (Schettini and Corchs, 2010). This paper proposes an algorithm based on the former where some of unknown parameters in the model are estimated for the given environment.

2.1. Jaffe–McGlamery underwater image model

Jaffe–McGlamery image model (Jaffe, 1990) has been the well-known complete radiometric model for the underwater image formation. The model is based on two basic assumptions: linear superposition of irradiance and attenuation modeling for medium-light interaction.

(1) *Linear superposition*: Jaffe–McGlamery model decomposes the irradiance at a specific point into three components: direct component, forward-scattered component, and backscatter component. These three components are linearly summed and they form the total irradiance as follows:

$$E_t = E_d + E_f + E_b \tag{1}$$

where E_t is the total irradiance, E_d the direct component, E_f the forward-scattered component, and E_b the backscatter. The direct component is the light reflected by the object surface and entered the camera without scattering and the forward-scattered component is the amount of light that is reflected by the object and entered the camera after scattered at a small angle. The backscatter is the light reflected not by the object but still entered the camera due to the scattering in the medium.

(2) *Attenuation modeling for medium-light interaction*: The other assumption for Jaffe–McGlamery model is that the intensity of the light traveling in a liquid decreases exponentially. As described in Yamashita et al. (2007), it is shown in the following equation:

$$L_i(z) = L_{0,i} \exp(-c_i z) \tag{2}$$

where i is the wavelength of light, z is the distance traveled, $L_i(z)$ is the light intensity of wavelength i, $L_{0,i}$ is the light intensity of

wavelength i at the light source, and c_i is the attenuation coefficient at wavelength i, respectively. As the light travels from the source to a point at distance z, the intensity decreases exponentially as a function of the attenuation coefficient and the traveled distance. Note that the attenuation coefficient is dependent on the wavelength of the light—the amount of attenuation varies as the wavelength changes.

Based on the two assumptions described above, the complete Jaffe–McGlamery underwater image model has been derived. Due to the scope of this paper, the full mathematical expression of the model is not shown. Refer to Jaffe (1990) for the complete presentation of the model.

2.2. Color restoration

Based on the Jaffe–McGlamery model described above, the color restoration algorithm has been developed with a number of further simplifications. Although the irradiance has been decomposed into three components, for the purpose of the color recovery method proposed in this paper, the focus has been given to the direct component only. It has been verified by Zhang and Negahdaripour (2002) that the backscatter component does not carry information about the scene. And the effect of the forward-scattered component can be treated as blurring. Therefore, the proposed restoration algorithm only focuses on the direct component of the irradiance with the consideration of intensity attenuation.

Fig. 1 shows the light traveling from the light source, reflected on the object surface and entering the camera. This is when considering only the direct component of the light. As demonstrated in Yamashita et al. (2007), if there is no attenuation while traveling, the intensity of the light can be modeled as follows:

$$I_i(l,z) = \frac{L_{0,i}\kappa_i \cos^3 \alpha}{z^2} \tag{3}$$

where l is the distance between the object surface and the camera, z is the distance between the source and the object, $I_i(l,z)$ is the light intensity of wavelength i, κ_i is the reflectance ratio of the objects surface of wavelength i, and α is the angle between the ray vector of light from the source and the normal vector of objects surface, respectively. This is the case when air is the traveling medium. For the case where the light travels in a liquid, the second assumption of the Jaffe–McGlamery model can be applied to the above equation as follows (Yamashita et al., 2007):

$$I_i(l,z) = \frac{L_{0,i}\kappa_i \cos^3 \alpha}{z^2} \exp\left\{-c_i\left(\frac{z}{\cos\alpha} + \frac{l}{\cos\theta}\right)\right\} \tag{4}$$

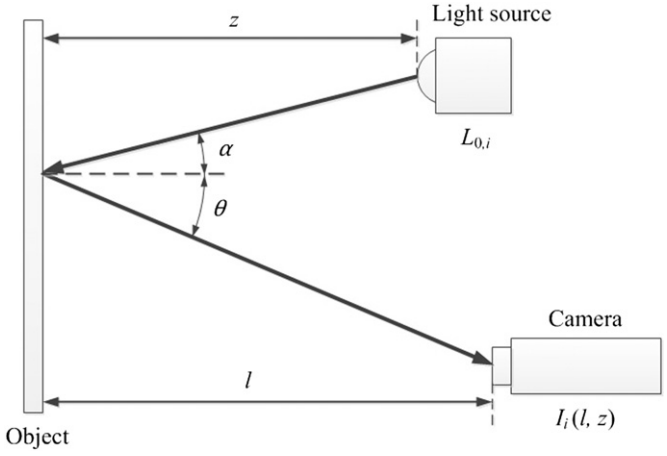

Fig. 1. Light refraction in liquid.

where θ is the angle between the ray vector of light from the object and the optical axis of lens. This expression shows the intensity of the light attenuated exponentially as a function of the distance traveled. However, the parameters such as light intensity at the source or the reflectance coefficient are unknown and difficult to find in real situations.

The main objective of the proposed color restoration method is to recover the degraded underwater color information so that the recovered images are similar to those taken in the air. It is necessary to define the relationship between the two intensity values in two different mediums: air and water. From the above equation, the relationship can be expressed as follows (Yamashita et al., 2007):

$$I_i(l',z',c_i') = I_i(l,z,c_i)\left(\frac{z}{z'}\right)^2 \frac{\cos^3 \alpha'}{\cos^3 \alpha} \times \exp\left\{-c_i'\left(\frac{z'}{\cos\alpha'} + \frac{l'}{\cos\theta'}\right) + c_i\left(\frac{z}{\cos\alpha} + \frac{l}{\cos\theta}\right)\right\} \tag{5}$$

where $I_i(l',z',c_i')$ and $I_i(l,z,c_i)$ denote the intensity of wavelength i in air and water, respectively. Other notations remain the same as in (Fig. 1) where the primes ($'$) represent the values in air. As denoted in Yamashita et al. (2007), this expression does not contain the reflectance ratio or the intensity of light at the source. With the equation above, the color information can be recovered using a number of geometric parameters and the attenuation coefficients for the two mediums. Therefore, the color restoration becomes the problem of estimating these values.

3. Object detection and tracking

3.1. Design of underwater target objects

Although it is desirable that the underwater robots have the ability to explore the natural unstructured environments with full autonomy, the depth of the research in the field has not reached this level. Moreover, due to the harsh conditions in underwater environments, the research in the field of vision-based underwater navigation is still in the early stages. For the purpose of this study, artificial underwater target objects have been designed, which are described in this section.

Regarding the missions for the underwater robots, many cases involve visual detection and tracking of certain targets. Some of the examples include visual inspections of manmade objects, environmental monitoring, and underwater mine detections. In these applications, vision-based techniques for detecting the interest objects that might or might not be known prior to the situation are required. The main challenge is that the appearances of the objects in water are different from those in air. As mentioned earlier, this is one of the problems to be solved in this study. A number of 3D objects have been designed and they have been used as underwater targets in order to test and verify the proposed algorithms in this paper.

The target objects have been designed so that the changes in their appearances are small as the view point changes. In other words, although they are 3D objects, they appear almost same for horizontal movement of the robot. The use of these primitive targets simplifies the detection problem while allowing in-depth study on the underwater image properties such as color degradation. The designed target objects are shown in Fig. 2. The robot's horizontal movement yields no significant changes in the captured image of the target objects due to their geometric symmetry. Also by restricting the vertical movement very small, we can safely assume little variation for view point changes. The color of the objects has been chosen as red since its observability in the

underwater environment is high and at the same time the attenuation of this color is not so large as distance increases.

3.2. Target object detection

As target objects in underwater environments vary in different applications, the detection of the interest target objects becomes an object detection and recognition problem. A number of recent computer vision techniques have been applied to the designed target objects and their feasibilities in underwater images have been analyzed.

3.2.1. Feature-based approach

One of the methods for object detection is to analyze the correspondences between the feature points detected in two different images. By looking at the similarities of the feature descriptors and their spatial distributions, decisions can be made for the object recognition. There has been a long history of the researchers trying to find the optimal ways of detecting, describing, and matching the interest image features.

Scale-invariant feature transform (SIFT) proposed by Lowe (2004) is known as the most reliable feature detector/descriptor. SIFT algorithm includes following four steps:

(i) scale-space extrema detection,
(ii) keypoint localization,
(iii) orientation assignment,
(iv) keypoint descriptor.

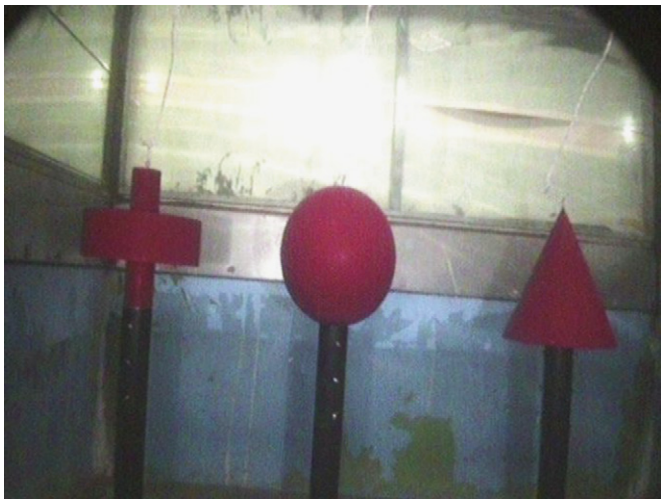

Fig. 2. Underwater target objects.

In the *scale-space extrema detection* step, the candidates for the interest points are detected where they are the strong survivors in the scale change. Once the candidates are defined, their exact locations are calculated in *keypoint localization* step. In the *orientation assignment* step, for each of the key points, its unique orientation is assigned based on the intensity gradients around the key point. And along with the assigned orientations, the descriptors are generated using gradient vectors in *keypoint descriptor* step.

Speeded up robust features (SURF) can be considered as a simplified version of SIFT (Bay et al., 2006). In order to reduce the computation time of SIFT, two new concepts have been added in SURF. For the detection of interest points, box filters are used instead of calculating the Laplacians of Gaussian. For further optimization, it is possible to reduce the length of the descriptor vector to a half (64-vector) compared to the original length (128-vector) of SIFT.

As it is one of the state-of-the-art methods of feature-based image matching, SURF has been tested for our purpose of object detection. The images of target objects were given prior to the test and SURF key points were extracted and kept in the database. When the images containing some of the target objects were given as inputs, the key points in the input images were extracted and compared with the database in order to determine the presence and the location of the target objects. These feasibility tests were carried with the images taken in air. Fig. 3 shows some of the test results.

In Fig. 3, the target image patches are shown in the upper left corners and the input images are shown in the lower portions. Lines connecting the two images denote the matched pairs of two feature points and the boxes drawn in the input images show the successful detection of the objects. The left two pictures in the figure show successful detections of the two objects whereas the picture on the right shows a failure. The results were worse when there are more than one target objects seen in the input images. In this example, the success or failure in the detection was manually determined. But we can consider the conservation of the geometric relationship of the detected features for automatic discrimination.

The feature-based approach for the target object detection was not successful and it has been concluded that this approach was not suitable for the designed target objects. There are several reasons for the results. One of them is the simple and textureless object designs. The feature descriptors generated from the key points are extremely simple and distinctions among them are too small to judge the differences. Another reason is that the target objects have similar feature points where their descriptors are almost identical. Although the outer contours are clearly different from one another, they have similar features in the local key point level. Confusion among the

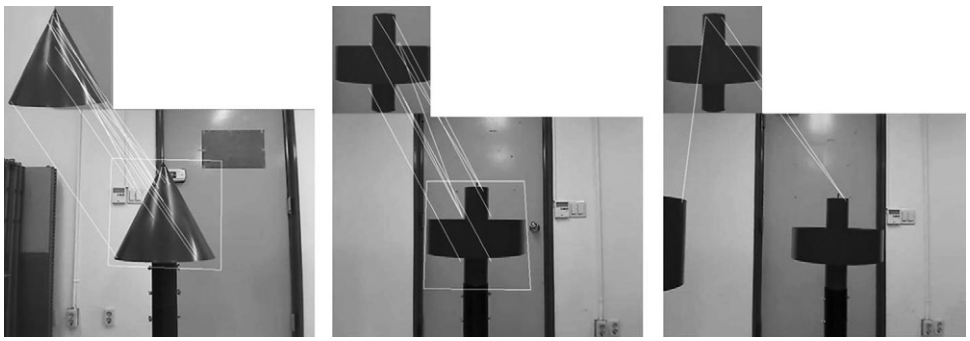

Fig. 3. Feature matching test using SURF in air. (Upper left corners) Target image patches. (Lower portions) Input images. White lines denote the matched pairs of two feature points. White boxes indicate the successful detection of the objects.

local features is directly related to the object recognition results in the feature-based object detection approach.

3.2.2. Template-based approach

Considering both the characteristics of the designed target objects and the underwater environments, object detection problem has been approached in another way while focusing on the contours of the objects. The template-based matching technique (Brunelli, 2009) can be used to localize the position of the interest objects in images based on the probabilistic approach. This method is more effective on the two-dimensional objects since it utilizes direct comparisons between the template and the input image.

Fig. 4 shows some of the feasibility test results using template matching technique. Each of the two pictures contains the template, input image with detection result, and the probabilistic map. Even with serious illumination changes and background changes, the detections were successful. The scale changes could be covered by changing the template sizes. This template-based approach focuses on the global shapes rather than the local features and this fact leads to the better results compared to the feature-based approach.

The template-based approach works well in small affine transformation, but it performs poorly for large affine. Since we have assumed a small variation for view point changes, we can also assume small affine transformation. And the normalized cross correlation (NCC)-based template matching shows robust results in varying illumination environments. We can also perform various preprocessing techniques like histogram normalization to increase the robustness for illumination changes.

3.3. Target object tracking

Once the target objects are detected using a detection method, the positions of the objects should be tracked in the subsequent frames. Two vision-based object tracking methods have been reviewed and their feasibilities in our application have been analyzed.

3.3.1. Optical flow tracking

One of the popular vision-based tracking methods is the optical flow tracking. Optical flow tracking estimates the positions of the image points by calculating the flow velocities in two consecutive frames. Lucas–Kanade optical flow method (Lucas and Kanade, 1981) is widely used where the computation load is decreased by implementing image windows.

Since the optical flow tracking requires specific image points to be tracked, the corner features have been extracted in the detected object areas for the tracking purpose. Tracking of the objects was successful when the number of reliable corner features was large, however, in some cases it was difficult to extract a good set of features. Not only that the number of stable features such as sharp corners was small, but also there were confusions among the features extracted on the textureless surfaces of the designed objects.

Some of the test results are shown in Fig. 5. Similar to the results of SURF feasibility test, the feature-based approach for the blurred underwater images was not successful. There were fewer feature points inside the cone- and sphere-shaped objects than the cross-shaped object. The reason is that the cross-shaped object has a little more complex shape than the others.

3.3.2. Mean shift tracking

Mean shift tracking is another vision-based tracking method based on the color histogram of the interest area. Mean shift is particularly effective for tracking non-rigid objects. Once the interest area, or the target object, is detected using one of the detection methods described above, the characteristics of the region is represented by a histogram. The color histogram is composed of 16 bins for each of the RGB channels and the full histogram is a vector with 4096 values. Once the histogram is generated, the shift of the mean of the histogram is calculated in the next frame in order to locate the position of the interest area. For detailed explanations, refer to Comaniciu et al. (2003).

Fig. 6 shows the mean shift algorithm applied to a sequence of underwater images. The initial detection was done using the template matching method described in Section 3.2 and the mean shift has been applied for the tracking. The three images show the successful tracking of three target objects where the yaw angle of the robot was changed clockwise.

Once the target objects were correctly located using the template matching method, the result of tracking afterwards was reliable. In addition, the mean shift tracking is able to track the scale change of the interest regions and therefore it was

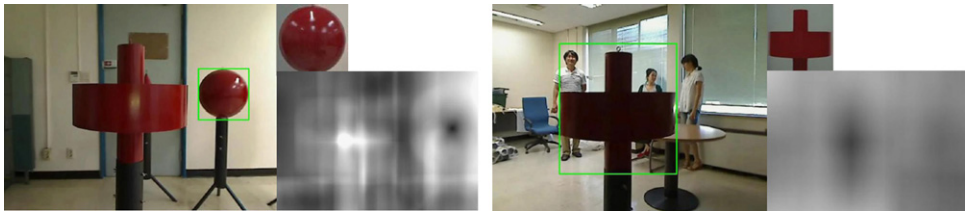

Fig. 4. Template matching test in air. (Upper right) Template images. (Left) Input images with detection result. (Lower right) Probabilistic correlation map.

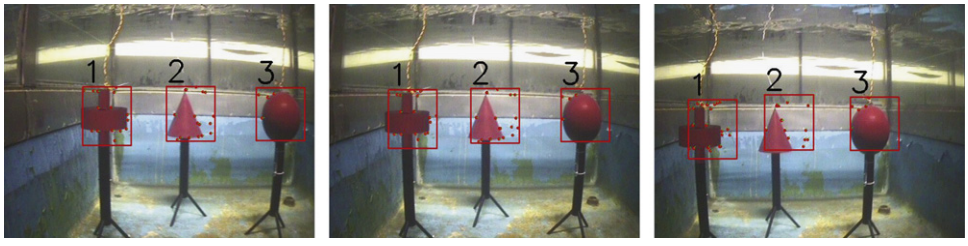

Fig. 5. Optical flow tracking. There are fewer feature points inside the cone and sphere shaped objects than the cross shaped object. Labels 1, 2, and 3 in the images denote cross, cone, and sphere, respectively.

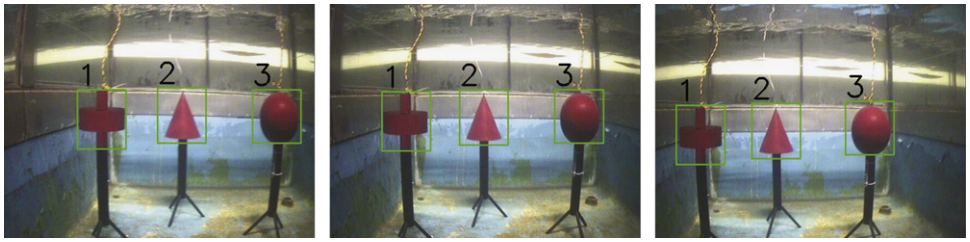

Fig. 6. Mean shift tracking. Three images show the tracking of three target objects where the yaw angle of the robot was changed clockwise. Labels 1, 2, and 3 in the images denote cross, cone, and sphere, respectively.

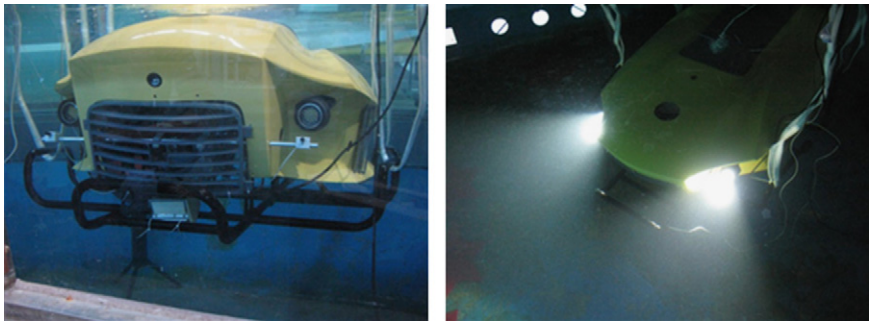

Fig. 7. Underwater robot platform *yShark*.

robust to the distance changes. On the other hand, it was not effective when two objects overlapped on top of each other in the image plane. Confusions occurred when comparing the histograms of the similar color properties.

4. Experimental results

The descriptions on the experiment setup and the results are discussed in this section. The topics include introduction of the robot platform used in the experiments, results of the color image restoration algorithm applied to the underwater images, and target detection and tracking results.

4.1. Robot platform

This section describes the details of the robot platform that has been used for the verification of the algorithm. We used an underwater robot named *yShark* which has been developed by Korea Ocean Research & Development Institute (KORDI). Main features of yShark shown in Fig. 7 are described here.

yShark is about 1.9 m long, 1.2 m wide, and weighs 70 kg. There are two pressure vessels; one for two lithium polymer battery packs, the other for a computer control system and sensor processing units where a computer controls the power and signal of all devices. It has four horizontal thrusters and two vertical thrusters to achieve fast target searching with depth keeping control. This thruster configuration provides five degrees of freedom without roll motion. Redundancy of horizontal thruster provides important functionalities such as accurate perception and fine motion. The control system consists of three PC104s (Lippert 1.8 GHz) that communicate through a high-speed internal network; a main control computer, an optical and sonar image processing computer, and an acoustic signal processing computer. There are two cameras with two LED lights. One is facing front and the other is facing downward. Bowtech Divecam-550C is used as a underwater camera. The cameras are enclosed with an acrylic housing rated to 100 m of operation. Sony 1/3" EX-View HAD CCD sensor offers high-resolution images with high-level sensitivity achieved by using 10 bit digital processing. Since this camera captures images up to 30 frames per second, the real-time implementation is possible.

4.2. Color image restoration

The degraded color qualities of underwater images can be recovered based on the algorithm described in Section 2. For parameter estimation and further simplification, the restoration Eq. (5) is revisited. The unknown geometric parameters that are dependent on the environments are the distances between the light sources and the objects (z and z'), distances between the object surfaces and the cameras (l and l'), the approaching angles of the ray vectors (α and α'), and the departing angles (θ and θ') for the two different mediums. Refer to Fig. 1 for the notations. These geometric parameters can be estimated according to the experimental settings. Considering the detectable ranges of the camera, location of the light source, and the geometric relationship between the robot and the objects, the parameters can be estimated for the further simplification of the equation. For the experiments carried out in this study, the following values have been used for the geometric parameters: $z=z'$, $l=l'$, $z\approx l$, $\theta=\theta'=0^\circ$, and $\alpha=\alpha'=0^\circ$.

This is a reasonable estimation of the parameters when considering the experimental settings. We assumed that the dominating light source in underwater is the light attached to the robot itself, thus the angle of the light ray becomes approximately zero. Considering the detection ranges and the field of the view of the camera, the distance was assumed to be constant. Although this simplification might degrade the restoration performance, it allows the real-time application of the detection algorithm. Moreover, these assumptions are not only valid for the experiments in this study, but also reasonable for the deep sea environments. In deep seas, these assumptions make more sense due to a number of facts such as there are no light sources other than the one from the robot and the detection range is limited in more turbid water.

Other unknowns remaining in Eq. (5) are the attenuation coefficients (c_i and c_i') of the traveling mediums. Attenuation coefficients are functions of the wavelengths of light. For the

three color channels, red, green, and blue, the attenuation coefficients have been set as in Table 1.

Attenuation of the light intensity occurs most at the higher frequencies and less in lower ones. In terms of the color channels, red attenuates most and green and blue are less affected. These values are the typical values, experimentally determined, for the shallow water conditions (Yamashita et al., 2007). On the other hand, it has been assumed that there are no light attenuations in air.

With the estimation of the geometric parameters and the attenuation coefficients as described above, the equation for the color restoration for each color channel becomes as follows:

$$I_{600}(d,d,0.0) = I_{600}(d,d,0.10)e^{0.20} \quad (6)$$

Table 1
Attenuation coefficients for RGB color channels.

Color channels	Wavelength (i)	Attenuation coefficient (c_i)
Red	600	0.10
Green	550	0.02
Blue	450	0.02

$$I_{550}(d,d,0.0) = I_{550}(d,d,0.02)e^{0.04} \quad (7)$$

$$I_{450}(d,d,0.0) = I_{450}(d,d,0.02)e^{0.04} \quad (8)$$

where d is the estimated distance to the object.

Now the equations above allow the simplified recovery of the degraded underwater colors. In the implementation point of view, these equations adjust the intensity magnitudes of the RGB channels with exponential functions. This is similar to the color mapping techniques used for color correction purposes. Developed from the basic model of the underwater images, the described algorithm makes the color restoration as a simple color mapping problem. Although the full restoration models with no simplifications can recover the degradation more accurately, this simplified version still enhances the color qualities for the designated environments while being able to process the images in real time.

Fig. 8 shows the images before and after the restoration. Due to the reasonably clean water condition and the short distances ($d \approx 1$ m) to the objects, the changes from the original images are not visually noticeable. In quantitative point of view, the changes in the pixel intensities after the restoration process are summarized in Table 2.

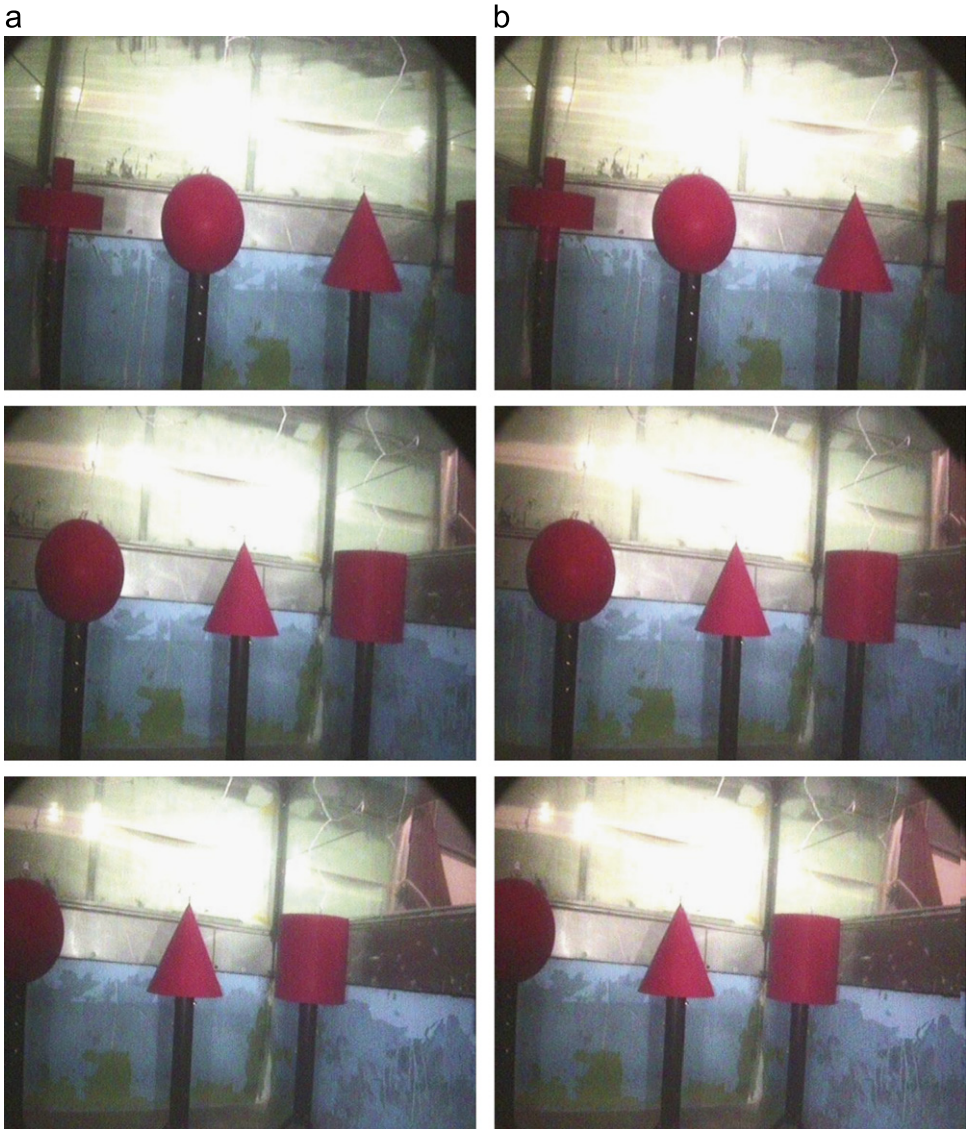

Fig. 8. Color restoration results: (a) raw and (b) restored images. (For interpretation of the references to color in this figure caption, the reader is referred to the web version of this article.)

The intensity of red has been increased by approximately 50% and there have been little changes in green and blue. The effectiveness of this restoration process on the target detection is discussed in the next section.

4.3. Target object detection results

The experiments for the tracking of the target objects have been carried out in an indoor test pool at Korea Ocean Research and Development Institute (KORDI) in Daejeon, Korea. The video clips taken by the underwater camera (Bowtech Divecam-550c) have been used for the evaluation of the detection algorithm.

The template matching algorithm has been implemented for the detection of the designed target objects. In addition to the detection algorithm tests, in order to examine the effectiveness of the image restoration algorithm on the detection results, a set of further experiments have been carried out. In more detail, the detection results with and without the camera calibration and the color restoration have been compared. And two different sets of the target templates, taken in air and in water respectively, have been tried for the template matching. The entire set of experiments is designed as in Table 3 for the reasons as follows.

For the detection purpose, it is most desirable to use the sample images of the target objects gathered in the actual underwater environment. The appearances and the color presentations of the objects in the same conditions make the detection easier. This is the case of Experiment 1. However, it is difficult to obtain the in-water templates in most of the real environmental situations. In more practical and realistic cases, the images of the target objects taken in air might be given prior to the actual underwater missions. This is the case of Experiments 2 and 3. For Experiment 3, the proposed color restoration algorithm has been applied in order to examine its effectiveness.

For the performance criteria, TPR (True Positive Rate) and FPR (False Positive Rate) have been used. TPR, or Sensitivity, represents the ratio of the true positives (TP) to the actual trues (TPR=TP/(TP+FN)). FN stands for the false negative. FPR, or 1-Specificity, can be calculated as FPR=FP/(FP+TN). FP and TN stand for the false positive and the true negative, respectively. TPR is the proportion of correct detections of the target object that yield positive test outcomes. FRP is the rate of absent detection events that also yield positive test outcomes. Larger TPR and smaller FPR are desirable.

For the experiments, 1629 frames of the video clips were used. Detection results of the experiments are shown in Tables 4–6. As shown in three tables, the detection with the in-water template (Experiment 1) shows the best results with TPR of 0.7973 and FPR of 0. It is obvious since the exact color appearance in the underwater situation is given prior to the detection. In Experiment 2, where the in-air template has been used without color restoration, TPR is 0.5528 and FPR is 0.1971. Compared to Experiment 1, the performance is worse in both criteria. When the restoration algorithm is applied in Experiment 3, TPR is increased to 0.6562 and FPR is 0.2083 which is almost the same as Experiment 2. We can see the effectiveness of the color restoration algorithm by comparing the results from these three experiments.

4.4. Target object tracking results

The videos of the target objects have been recorded using the underwater camera as in the previous experiments. The motion of the vehicle has been controlled by the operator while recording.

The mean shift algorithm explained in Section 3.2 has been used for the object tracking purpose. Fig. 9 shows a series of images while tracking the four designed objects using mean shift. The mean shift tracking algorithm runs for each of the detected object. In other words, detection of a new object triggers a new tracker for the specific object. For example, when there are four

Table 2
Intensity changes after restoration.

Color channels	Intensity change factors
Red	1.22
Green	1.04
Blue	1.04

Table 3
Experiment conditions.

Exp. sets	Conditions	
	Color restoration	Templates
1	N	In-water
2	N	In-air
3	Y	In-air

Table 5
Detection results of Experiment 2.

Performance criteria	Cross	Cone	Sphere	Cylinder	Total
TP	165	470	865	568	2068
FN	445	835	63	330	1673
FP	0	0	1	546	547
TN	1019	324	700	185	2228
TPR (sensitivity)	0.2705	0.3602	0.9321	0.6325	0.5528
FPR (1-specificity)	0	0	0.0014	0.7469	0.1971

Table 6
Detection results of Experiment 3.

Performance criteria	Cross	Cone	Sphere	Cylinder	Total
TP	199	747	873	636	2455
FN	411	558	55	262	1286
FP	0	0	2	576	578
TN	1019	324	699	155	2197
TPR (sensitivity)	0.3262	0.5724	0.9407	0.7082	0.6562
FPR (1-specificity)	0	0	0.0029	0.7880	0.2083

Table 4
Detection results of Experiment 1.

Performance criteria	Cross	Cone	Sphere	Cylinder	Total
TP	474	1237	581	694	2986
FN	136	68	351	204	759
FP	0	0	0	0	0
TN	1019	324	697	731	2771
TPR (sensitivity)	0.7771	0.9479	0.6234	0.7728	0.7973
FPR (1-specificity)	0	0	0	0	0

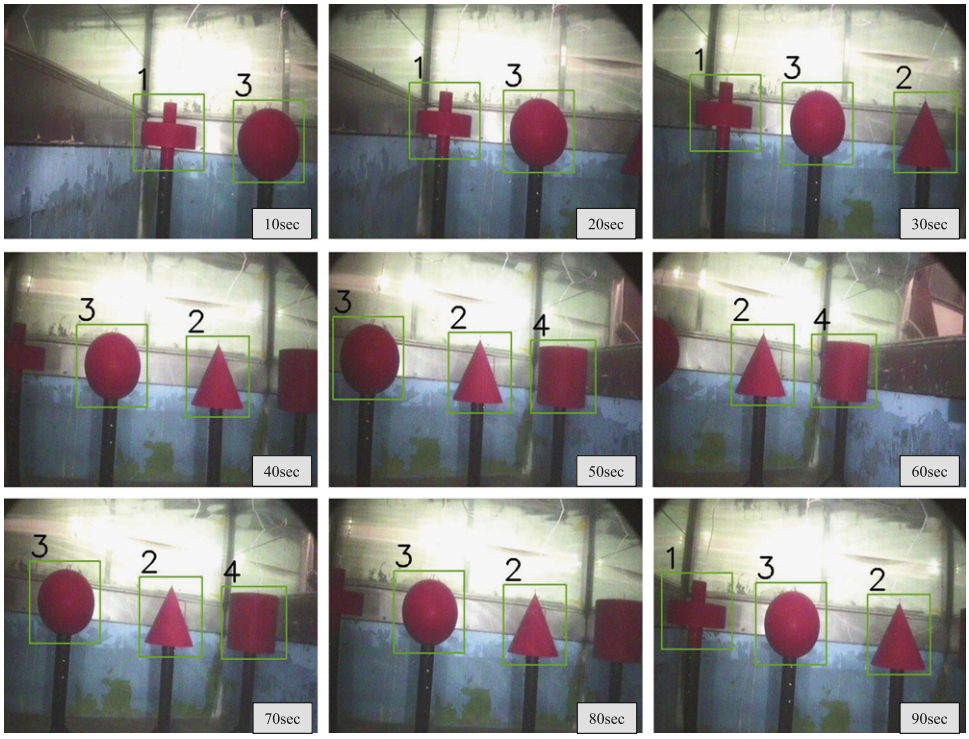

Fig. 9. Target object tracking using mean shift. Labels 1, 2, 3, and 4 in the images denote cross, cone, sphere, and cylinder, respectively

objects detected in a series of images, there are four mean shift trackers running at the same time. While tracking, when the scale of the object changes suddenly or the position shifts by a large amount, which are not likely to happen in underwater environments, the tracker is assumed to has been failed and the detection algorithm is performed for a new search. Since the proposed color restoration algorithm improves the performance of the object detection, the underwater object tracking experiment also shows better results.

5. Conclusions

In this work, vision-based techniques for autonomous navigations of underwater robots have been studied in depth. Due to the poor visibility in underwater and the limited detection ranges of cameras, vision has not been in the center of the underwater sensing and considered as less informative compared to other underwater sensors. In order to overcome these limitations and to make use of the full advantages of image data, a number of approaches have been taken in this study. The underwater color restoration algorithm has been developed from the full underwater radiometric model and it has been applied and tested in the real time image processing. Performance improvements from the restorations have been analyzed in the object detection algorithms. In addition, the detection and tracking of artificial targets using underwater images are studied using a number of computer vision techniques, and the experiments have been conducted using an underwater robot platform.

As the further improvement of the current study, the future work includes (i) conducting more qualitative and quantitative analysis on the results of the underwater image restoration algorithm and (ii) employing other forms of target objects for the detection and tracking purpose. Once fully tested with the artificial targets, the final goal is to develop a robust and autonomous vision-based navigation method in the natural underwater environments.

Acknowledgments

This research was supported by Korea Ocean Research and Development Institute (KORDI). (Project title: "Development of technologies for an underwater robot based on artificial intelligence for highly sophisticated missions").

References

Balasuriya, A., Ura, T., October 2002. Vision-based underwater cable detection and following using AUVs. In: Proceedings of the MTS/IEEE OCEANS 2002, vol. 3, pp. 1582–1587.

Balasuriya, B., Takai, M., Lam, W., Ura, T., Kuroda, Y., October 1997. Vision based autonomous underwater vehicle navigation: underwater cable tracking. In: Proceedings of the MTS/IEEE OCEANS 1997, vol. 2, pp. 1418–1424.

Bay, H., Tuytelaars, T., Gool, L.V., 2006. SURF: speeded up robust features. In: Proceedings of the European Conference on Computer Vision, vol. 3951, pp. 404–417.

Brunelli, R., 2009. Template Matching Techniques in Computer Vision: Theory and Practice. Wiley.

Caffaz, A., Caiti, A., Casalino, G., Turetta, A., 2010. The hybrid glider/AUV folaga. IEEE Robotics Automation Magazine 17 (March (1)), 31–44.

Comaniciu, D., Ramesh, V., Meer, P., 2003. Kernel-based object tracking. IEEE Transactions on Pattern Analysis and Machine Intelligence 25 (May (5)), 564–577.

Corke, P., Detweiler, C., Dunbabin, M., Hamilton, M., Rus, D., Vasilescu, I., April 2007. Experiments with underwater robot localization and tracking. In: Proceedings of the IEEE International Conference on Robotics and Automation 2007, pp. 4556–4561.

Dudek, G., Jenkin, M., Prahacs, C., Hogue, A., Sattar, J., Giguere, P., German, A., Liu, H., Saunderson, S., Ripsman, A., Simhon, S., Torres, L., Milios, E., Zhang, P., Rekletis, I., August 2005. A visually guided swimming robot. In: Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems, 2005, pp. 3604–3609.

Feezor, M.D., Sorrell, F.Y., Blankinship, P.R., Bellingham, J.G., 2001. Autonomous underwater vehicle homing/docking via electromagnetic guidance. IEEE Journal of Oceanic Engineering 26 (October (4)), 515–521.

Jaffe, J., 1990. Computer modeling and the design of optimal underwater imaging systems. IEEE Journal of Oceanic Engineering 15 (April (2)), 101–111.

Kim, A., Eustice, R., October 2009a. Pose-graph visual SLAM with geometric model selection for autonomous underwater ship hull inspection. In: IEEE/RSJ International Conference on Intelligent Robots and Systems 2009, pp. 1559–1565.

Kim, A., Eustice, R., October 2009b. Toward AUV survey design for optimal coverage and localization using the Cramer Rao Lower Bound. In: MTS/IEEE OCEANS 2009, pp. 1–7.

Lee, P.-M., Jeon, B.-H., Kim, S.-M., September 2003. Visual serving for underwater docking of an autonomous underwater vehicle with one camera. In: Proceedings of the OCEANS 2003, vol. 2, pp. 677–682.

Lowe, D.G., 2004. Distinctive image features from scale-invariant keypoints. International Journal on Computer Vision 60 (2), 91–110.

Lucas, B.D., Kanade, T., 1981. An iterative image registration technique with an application to stereo vision. In: Proceedings of the 1981 DARPA Imaging Understanding Workshop, pp. 121–130.

Maire, F., Prasser, D., Dunbabin, M., Dawson, M., 2009. A vision based target detection system for docking of an autonomous underwater vehicle. In: Proceedings of the 2009 Australasian Conference on Robotics and Automation.

Marani, G., Choi, S., 2010. Underwater target localization. IEEE Robotics Automation Magazine 17 (March (1)), 64–70.

Negre, A., Pradalier, C., Dunbabin, M., 2008. Robust vision-based underwater target identification and homing using self-similar landmarks. Journal of Field Robotics 25, 360–377.

Papanikolopoulos, N.P., Khosla, P.K., 1993. Adaptive robotic visual tracking: theory and experiments. IEEE Transactions on Automatic Control 38 (March (3)), 429–445.

Psarros, D., Papadimitriou, V., Chatzakos, P., Spais, V., Hrissagis, K., 2010. A service robot for subsea flexible risers. IEEE Robotics Automation Magazine 17 (March (1)), 55–63.

Sattar, J., Dudek, G., May 2009. Robust servo-control for underwater robots using banks of visual filters. In: Proceedings of the IEEE International Conference on Robotics and Automation 2009, pp. 3583–3588.

Schettini, R., Corchs, S., 2010. Underwater image processing: state of the art of restoration and image enhancement methods. EURASIP Journal on Advances in Signal Processing, 2010, 746025.

Trucco, E., Olmos-Antillon, A.T., 2006. Self-tuning underwater image restoration. IEEE Journal of Oceanic Engineering 31 (April (2)), 511–519.

Yamashita, A., Fujii, M., Kaneko, T., April 2007. Color registration of underwater images for underwater sensing with consideration of light attenuation. In: Proceedings of the IEEE International Conference on Robotics and Automation 2007, pp. 4570–4575.

Yu, S.-C., Ura, T., Fujii, T., Kondo, H., 2001. Navigation of autonomous underwater vehicles based on artificial underwater landmarks. In: Proceedings of the MTS/IEEE OCEANS 2001, vol. 1, pp. 409–416.

Zhang, S., Negahdaripour, S., 2002. 3-D shape recovery of planar and curved surfaces from shading cues in underwater images. IEEE Journal of Oceanic Engineering 27 (January (1)), 100–116.